

Efficient learning algorithm for quantum perceptron unitary weights

Kok-Leong Seow

Department of Electrical Engineering and Computer Science
Wichita State University
kxseow1@wichita.edu

Elizabeth C. Behrman

Department of Mathematics, Statistics, and Physics
Wichita State University
elizabeth.behrman@wichita.edu

James E. Steck

Department of Aerospace Engineering
Wichita State University
james.steck@wichita.edu

Abstract

For the past two decades, researchers have attempted to create a Quantum Neural Network (QNN) by combining the merits of quantum computing and neural computing. In order to exploit the advantages of the two prolific fields, the QNN must meet the non-trivial task of integrating the unitary dynamics of quantum computing and the dissipative dynamics of neural computing. At the core of quantum computing and neural computing lies the qubit and perceptron, respectively. We see that past implementations of the quantum perceptron model have failed to fuse the two elegantly. This was due to a slow learning rule and a disregard for the unitary requirement. In this paper, we present a quantum perceptron that can compute functions uncomputable by the classical perceptron while analytically solving for parameters and preserving the unitary and dissipative requirements.

Keywords Quantum computing, Artificial neural networks, Unitary, Quantum neural networks, Singular value decomposition

1 Introduction

The idea of quantum mechanics being computationally more powerful than a classical Turing machine was first proposed by Feynman in [1, 2]. Thus far, there have been only two algorithms that utilize the power of quantum computing: Shor's factorization [3] and Grover's database search [4]. As quantum computing technology evolves, it seems natural to combine the properties of artificial neural networks with quantum computing. The synergy of quantum and neural computing will grant us advantages unattainable by classical artificial networks.

As stated in [5], quantum neural networks have advantages in the following areas: 1) exponential memory capacity [6], 2) higher performance for lower number of hidden neurons [7], 3) faster learning [8], 4) elimination of catastrophic forgetting due to the absence of pattern interference [8], 5) single layer network solution of linear inseparable problems [8], 6) absence of wires [9], 7) processing speed [9], 8) small scale [9], and 9) higher stability and reliability [7]. While Kak introduced the concept of applying quantum mechanics to artificial neural networks [10], the first in depth examination of QNNs were provided in Menneer's PhD thesis [11]. Since then, a myriad of attempts to combine the strengths of quantum computing and neural computing have been presented, and are reviewed in [12].

In the review, Shuld *et al.* classified all QNN models into 4 different approaches: 1) interpreting the step-function as measurement in order to combine the non-linearity of neural networks and linearity of quantum computing, 2) using quantum circuits that attempt to model the dynamics of neural networks, 3) describing the basic component of NNs, the perceptron, with a quantum formalism, and 4) modeling NNs with interacting quantum dots consisting of four atoms sharing two electrons. In this paper, we focus on the third approach of creating a quantum equivalent for a perceptron, with the goal to construct the building block for a more complex quantum neural network.

The first attempt to develop a quantum perceptron to overcome the limitations imposed by classical perceptrons was in 2001 by Altaisky [13]. Altaisky implemented his quantum perceptron through the use of optical modes, optical beam splitters, and phase shifters. Altaisky directly translated the perceptron model and replaced it with $|y\rangle = \hat{F} \sum_{j=1}^N \hat{w}_j |x_j\rangle$ along with the quantum equivalent of the learning rule,

$$\hat{w}_j(t+1) = \hat{w}_j(t) + \eta(|d\rangle - |y(t)\rangle) \langle x_j|. \quad (1)$$

Following Altaisky was Fei *et al.* [14], who implemented a new version of the iterative rule that could compute the XOR-function using one neuron, a feat not capable by classical perceptrons. Fei's activation operator and weights were unitary, but provided no way to derive the operators/weights which seemed specific to each problem. Next, Zhou *et al.* [15] also computed the XOR-function with one neuron, but required many iterations. Its learning rule implied a possible non unitary gate, contradicting the requirements of quantum computing. Slightly different from the previous models, Siomau introduced an autonomous quantum perceptron based on calculating a set of positive valued operators and value measurements (POVM) [16]. However, his perceptron could not replicate basic quantum gates such as quantum-Not gate and Hadamard gate, which are staples of quantum computing. Finally, Sagheer and Zidan [17] implemented their Autonomous Quantum Perceptron Neural Network by calculating the set of activation operators using the form: $\begin{pmatrix} \cos \theta_j & -\sin \theta_j \\ \sin \phi_j & \cos \phi_j \end{pmatrix}$. Similar to the aforementioned models, it also has an iterative method to learn the weights, as well as the issue of the learned weights/operators being non unitary. This severe violation creates a need for our proposed quantum perceptron that removes the iterative learning and analytically solves for weights/operators while maintaining the unitary requirements.

This paper is organized as follows: Section 2 describes the background mathematics. Topics included are the basics of quantum computing, neural networks, and singular value decomposition (SVD). Section 3 presents the the quantum perceptron and learning algorithm. We illustrate the perceptron's robustness by first applying the basic learning rule to the NOT gate and Hadamard gate to show that the newfound weights/operators took several iterations and are not unitary, which thus violates a strict requirement. To combat this issue, we present a closed-form solution to find the proper weights by using Moore-Penrose pseudoinverse and a tensor product. We illustrate that this also generalizes to two qubits and is capable of computing the controlled NOT function. We observe that in the case of computing our nonlinear controlled swap gate, that the unitary isn't preserved. We fix the issue through the use of SVD, and illustrate it's significance by computing the XOR function. We conclude in section 4.

2 Background

2.1 Quantum Computing

The basic elements of quantum computing are qubits, complex unit vectors in a 2-dimensional Hilbert space \mathcal{H}_2 . Labeled $|0\rangle$ and $|1\rangle$, the two quantum states express one bit of information: $|0\rangle$ corresponds to the bit 0 of classical computers, and $|1\rangle$ to the bit 1. More formally, given the basis $\{|0\rangle, |1\rangle\}$, a basic qubit is given in Dirac notation and described by the wavefunction

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad \alpha, \beta \in \mathbb{C} \quad (2)$$

where α and β are complex numbers called probability amplitudes that satisfy $|\alpha|^2 + |\beta|^2 = 1$. The state ψ of a general quantum system is in a linear superposition of basis states $|0\rangle$ and $|1\rangle$ with $|\alpha|^2$ probability of

collapsing into state $|0\rangle$ and $|\beta|^2$ probability of collapsing into state $|1\rangle$. The qubits may also be written in vector notation, $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

2.2 Neural Networks

Neurons are symbolized by variables $x, y = \{-1, 1\}$ where '1' indicates that the neuron is firing and '-1' indicates that it is resting. The activation mechanism of a neuron y due to the input of N other neurons forms the core of neural computing. This setup creates the simplest neural network possible: the perceptron. The perceptron is a computational model of a single neuron and is shown in figure ???. Given n input nodes called neurons with values $x_k = \{-1, 1\}, k = 1 \dots n$ that feed signals into a single output neuron y . Each input neuron is connected to the output neuron with a certain strength denoted by weight parameter $w_k \in [-1, 1]$. Neuron y is activated in an 'integrate-and-fire' mechanism, meaning that the post-synaptic signals are simply added up and compared with the specific threshold θ_y of neuron y . The input-output relation is governed by the activation function

$$y = \begin{cases} 1, & \text{if } \sum_{k=1}^n w_k x_k \geq \theta_y \\ -1 & \text{else} \end{cases} \quad (3)$$

The perceptron learning rule is governed by the equation

$$w_k(t+1) = w_k(t) + \eta(d - y)x \quad (4)$$

Worth mentioning, the classical perceptron is unable to calculate the non-separable XOR function.

2.3 Singular Value Decomposition

Any rectangular matrix A can be decomposed into three matrices defined as

$$A = U \Sigma V^T \quad (5)$$

where $U \in \mathbb{R}^{N \times N}$ is a unitary matrix, i.e., $U^T U = U U^T = I$ (identity matrix), $\Sigma \in \mathbb{R}^{N \times N}$ is diagonal matrix of singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, and $V \in \mathbb{R}^{M \times M}$ is a unitary matrix $V^T V = V V^T = I$ (identity matrix).

3 Quantum Neural Network

To describe the efficiency of our proposed quantum neuron and algorithm, we first present the quantum learning rule given in Eq. 1 to see where issues arise. We then counter the iterative method by analytically solving for weights by using Moore-Penrose pseudoinverse, a similar technique used by Extreme Learning Machine (ELM) [18] in the classical domain. Finally, to preserve the quantum properties of being unitary, we decompose our non-unitary matrix \hat{w} into three unitary matrices using our version of SVD.

3.1 Moore-Penrose Pseudoinverse

Given a quantum perceptron of the form $|y_{output}\rangle = \hat{F} \sum_{j=1}^N \hat{w}_j |x_j\rangle$, we can set $\hat{F} = I$ and obtain $|y_{output}\rangle = \sum_{j=1}^N \hat{w}_j |x_j\rangle$. We can then find the Moore-Penrose pseudoinverse of x , and analytically calculate \hat{w}_j through the equation

$$\hat{w}_j = |y_j\rangle \otimes |x_j^\dagger\rangle \quad (6)$$

where we take the tensor product of $|y_j\rangle$ and $|x_j^\dagger\rangle$. We first apply equation 1 to the NOT gate and Hadamard gate to show the possibility of non-unitary weights and slow convergence.

3.1.1 The NOT gate

Given two training examples

$$\begin{aligned} |x_1\rangle &= |0\rangle, \quad |y_1\rangle = |1\rangle \\ |x_2\rangle &= |1\rangle, \quad |y_2\rangle = |0\rangle \end{aligned}$$

utilizing the learning rule, we set $\hat{F} = I$ and randomize the initial weights to $w_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\eta = 0.1$, and target output $|y\rangle = |0\rangle$. Applying equation 1, we get

$$\begin{aligned} \hat{w}(t+1) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + 0.1(|0\rangle + |1\rangle)\langle 1| \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + 0.1(|0\rangle\langle 1| + |1\rangle\langle 1|) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 0.1 \\ 0 & 0.1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0.1 \\ 0 & -0.9 \end{pmatrix} \end{aligned}$$

after one iteration. We see that after 25 iterations we obtain

$$\begin{aligned} \hat{w} &= \begin{pmatrix} 1 & 0.9282102 \\ 0 & -0.0717898 \end{pmatrix} \\ \text{and } |y_{output}\rangle &= \begin{pmatrix} 1 & 0.9282102 \\ 0 & -0.0717898 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.9282102 \\ -0.0717898 \end{pmatrix}. \end{aligned}$$

It is easy to see that the calculated weights are not unitary, violating an important property of quantum computing. Instead, we calculate the Moore-Penrose pseudoinverse of the input and determine weights analytically. We apply equation Eq. 6 to inputs $|x_1\rangle$ and $|x_2\rangle$.

$$\begin{aligned} |x_1^\dagger\rangle &= \begin{pmatrix} 1 & 0 \end{pmatrix} \\ \hat{w}_1 &= |y_1\rangle \otimes |x_1^\dagger\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ |x_2^\dagger\rangle &= \begin{pmatrix} 0 & 1 \end{pmatrix} \\ \hat{w}_2 &= |y_2\rangle \otimes |x_2^\dagger\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ \hat{w} &= \sum_{j=1}^N \hat{w}_j = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

. Using our new weights, we compute our output.

$$\begin{aligned} |y_{output}\rangle &= \sum_{j=1}^N \hat{w} |x_1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \\ |y_{output}\rangle &= \sum_{j=1}^N \hat{w} |x_2\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle \end{aligned}$$

It is easy to see that our newly learned weights are indeed unitary, and after only one iteration.

3.1.2 The Hadamard gate

Given two training examples

$$\begin{aligned} |x_1\rangle &= |0\rangle, \quad |y_1\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \\ |x_2\rangle &= |1\rangle, \quad |y_2\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle \end{aligned}$$

by utilizing the learning rule, we set $\hat{F} = I$ and randomize the initial weights to $w_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\eta = 0.1$, and target output $|y\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. We see that after 25 iterations we obtain

$$\begin{aligned} \hat{w} &= \begin{pmatrix} 0.72813353 & 0 \\ 0.65634373 & -1 \end{pmatrix} \\ \text{and } |y_{\text{output}}\rangle &= \begin{pmatrix} 0.72813353 & 0 \\ 0.65634373 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.72813353 \\ 0.65634373 \end{pmatrix}. \end{aligned}$$

A quick observation shows us that these weights also violate the unitary condition. We again calculate the Moore-Penrose pseudoinverse of the input and determine weights analytically. We apply Eq. 6 to inputs $|x_1\rangle$ and $|x_2\rangle$.

$$\begin{aligned} |x_1^\dagger\rangle &= \begin{pmatrix} 1 & 0 \end{pmatrix} \\ \hat{w}_1 &= |y_1\rangle \otimes |x_1^\dagger\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \\ |x_2^\dagger\rangle &= \begin{pmatrix} 0 & 1 \end{pmatrix} \\ \hat{w}_2 &= |y_2\rangle \otimes |x_2^\dagger\rangle = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ \hat{w} &= \sum_{j=1}^N \hat{w}_j = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \end{aligned}$$

Using our computed weights, we determine our output

$$\begin{aligned} |y_{\text{output}}\rangle &= \hat{F} \sum_{j=1}^N \hat{w}_j |x_1\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \\ |y_{\text{output}}\rangle &= \hat{F} \sum_{j=1}^N \hat{w}_j |x_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle \end{aligned}$$

It's easily verifiable that our determined weights are indeed unitary, and after only one iteration.

3.1.3 The Controlled NOT gate

We show that the Moore-Penrose pseudoinverse is also applicable to 2-qubit systems. Given four training examples

$$\begin{aligned} |x_1\rangle &= |00\rangle, \quad |y_1\rangle = |00\rangle \\ |x_2\rangle &= |01\rangle, \quad |y_2\rangle = |01\rangle \\ |x_3\rangle &= |10\rangle, \quad |y_3\rangle = |11\rangle \\ |x_4\rangle &= |11\rangle, \quad |y_4\rangle = |10\rangle, \end{aligned}$$

after setting $\hat{F} = I$, we then calculate the Moore-Penrose pseudoinverse.

$$\begin{aligned}
|x_1^\dagger\rangle &= (1 \ 0 \ 0 \ 0) \\
|x_2^\dagger\rangle &= (0 \ 1 \ 0 \ 0) \\
|x_3^\dagger\rangle &= (0 \ 0 \ 1 \ 0) \\
|x_4^\dagger\rangle &= (0 \ 0 \ 0 \ 1)
\end{aligned} \tag{7}$$

We then take the tensor product of the pseudoinverse and input to create its corresponding weight. Summing up of all the individual weights, we acquire the final weight.

$$\begin{aligned}
\hat{w}_1 = |y_1\rangle \otimes |x_1^\dagger\rangle &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \hat{w}_2 = |y_2\rangle \otimes |x_2^\dagger\rangle &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\hat{w}_3 = |y_3\rangle \otimes |x_3^\dagger\rangle &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \hat{w}_4 = |y_4\rangle \otimes |x_4^\dagger\rangle &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\hat{w} = \sum_{j=1}^N \hat{w}_j &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}
\end{aligned}$$

By observation, it is easy to see that our algorithm generalizes to the 2-qubit system and are indeed unitary.

3.1.4 The Controlled Swap gate

We now test the dissipative features of our qubit by attempting to calculate the nonlinear version of the controlled swap gate. If qubit one is HIGH then swap qubits 2 and 3. Our eight states are shown below:

$$\begin{aligned}
|x_1\rangle &= |000\rangle, |y_1\rangle = |00\rangle, |x_2\rangle = |001\rangle, |y_2\rangle = |01\rangle \\
|x_3\rangle &= |010\rangle, |y_3\rangle = |10\rangle, |x_4\rangle = |011\rangle, |y_4\rangle = |11\rangle \\
|x_5\rangle &= |100\rangle, |y_5\rangle = |00\rangle, |x_6\rangle = |101\rangle, |y_6\rangle = |10\rangle \\
|x_7\rangle &= |110\rangle, |y_7\rangle = |01\rangle, |x_8\rangle = |111\rangle, |y_8\rangle = |11\rangle.
\end{aligned}$$

After calculating the weights by searching for the Moore-Penrose pseudoinverse and tensoring, we sum up all of the individual weights.

$$\hat{w} = \sum_{j=1}^N \hat{w}_j = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We see that in the nonlinear case, the weight matrix isn't unitary. In order to combat this, we introduce Singular Value Decomposition (SVD). SVD returns two unitary matrices and a diagonal matrix of singular values. We observe that there isn't a loss of generality by removing the diagonal matrix and replacing it with a unitary matrix of ones in the diagonals and zeros everywhere else. By setting

$$\hat{F}\Sigma_{new}\hat{w}_{new} = U\Sigma V^T \tag{8}$$

we can thus decompose the \hat{w} into three matrices. To improve accuracy we can also introduce a measurement function $M(x)$ such that

$$M = \begin{cases} 1, & \text{if } x > 0.5 \\ 0 & \text{else} \end{cases} \tag{9}$$

Applying SVD to \hat{w} we obtain

$$\hat{F} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Sigma_{new} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$w_{new}^{\wedge} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Calculating

$$M(\hat{F}\Sigma_{new}w_{new}^{\wedge}|x\rangle), \quad (10)$$

we consequently get the proper outcome in a closed form solution while maintaining unitary weights.

4 Proposed Algorithm

In this section we summarize the full algorithm and demonstrate its prowess by computing the XOR function. Given N data sets:

1. Set $\hat{F} = I$ (Identity matrix).
2. Calculate the Moore-Penrose pseudoinverse of each input.
3. Take the tensor product of the pseudoinverse of each input and its respective target to obtain the corresponding weight using Eq. 6.
4. Sum up all of the weights to create weight matrix \hat{w}
5. If \hat{w} is not unitary, move to step 6, else move to step 8.
6. Use SVD to decompose the \hat{w} into three matrices $\hat{F}\Sigma\hat{w}_{new}$
7. Pull out the diagonal matrix of singular values and replace it with a unitary matrix Σ_{new} with ones in the diagonals and zeroes elsewhere.
8. Use Eq. 10 to calculate $|y_{output}\rangle$.

4.0.5 The XOR function

The XOR function is nonlinear and uncomputable by a single neuron in the classical perceptron. Given a set of four different inputs and targets

$$\begin{aligned} |x_1\rangle &= |00\rangle, & |y_1\rangle &= |0\rangle \\ |x_2\rangle &= |01\rangle, & |y_2\rangle &= |1\rangle, \\ |x_3\rangle &= |10\rangle, & |y_3\rangle &= |1\rangle, \\ |x_4\rangle &= |11\rangle, & |y_4\rangle &= |0\rangle, \end{aligned}$$

we can place the inputs in a superposition governed by the wave function $|\psi\rangle$.

$$|\psi\rangle = \frac{1}{\sqrt{4}}|00\rangle + \frac{1}{\sqrt{4}}|01\rangle + \frac{1}{\sqrt{4}}|10\rangle + \frac{1}{\sqrt{4}}|11\rangle$$

After garnering the pseudoinverse of each state and tensoring with its respective target, we sum up all of the individual weights to obtain out final weight.

$$\hat{w} = \sum_{j=1}^N \hat{w}_j = \begin{pmatrix} \frac{1}{\sqrt{4}} & 0 & 0 & \frac{1}{\sqrt{4}} \\ 0 & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & 0 \end{pmatrix}$$

Decomposing \hat{w} into three unitary matrices, we now acquire

$$\hat{F} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Sigma_{new} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$w_{new}^{\hat{}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Calculating $M(\hat{F}\Sigma_{new}w_{new}^{\hat{}}|x\rangle)$ returns the proper output.

5 Conclusion

To harness the computational power of quantum computing and artificial neural networks, fundamental rules that grant them their intrinsic properties must not be infringed. In addition, QNNs should be robust enough to not be overly sensitive to the choice of operators as well as the ability to analytically solve for its parameters, similar to its classical counterpart. The work of the quantum perceptron model for the past two decades have failed to demonstrate these imperative qualities. In our proposed quantum perceptron and algorithm, we expel the iterative learning rule by analytically calculating parameters while maintaining unitary requirements. We also demonstrate its nonlinear capabilities by computing the XOR function, a feat impossible for the classical perceptron.

References

- [1] R. P. Feynman, “Quantum mechanical computers,” *Foundations of physics*, vol. 16, no. 6, pp. 507–531, 1986.
- [2] R. P. Feynman, “Simulating physics with computers,” *International journal of theoretical physics*, vol. 21, no. 6/7, pp. 467–488, 1982.
- [3] P. W. Shor, “Algorithms for quantum computation: Discrete logarithms and factoring,” in *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pp. 124–134, IEEE, 1994.
- [4] L. K. Grover, “A fast quantum mechanical algorithm for database search,” in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212–219, ACM, 1996.
- [5] A. A. Ezhov and D. Ventura, “Quantum neural networks,” in *Future directions for intelligent systems and information sciences*, pp. 213–235, Springer, 2000.
- [6] D. Ventura and T. Martinez, “Quantum associative memory,” *Information Sciences*, vol. 124, no. 1, pp. 273–296, 2000.
- [7] D. Cutting, “Would quantum neural networks be subject to the decidability constraints of the church-turing thesis,” *Neural Network World*, vol. 9, pp. 163–168, 1999.
- [8] T. Menneer and A. Narayanan, “Quantum-inspired neural networks,” *Department of Computer Science, University of Exeter, Exeter, United Kingdom, Technical Report*, vol. 329, p. 1995, 1995.
- [9] E. C. Behrman, J. E. Steck, and S. R. Skinner, “A spatial quantum neural computer,” in *Neural Networks, 1999. IJCNN’99. International Joint Conference on*, vol. 2, pp. 874–877, IEEE, 1999.
- [10] S. Kak, “On quantum neural computing,” *Information Sciences*, vol. 83, no. 3, pp. 143–160, 1995.
- [11] T. S. I. Menneer, *Quantum artificial neural networks*. PhD thesis, University of Exeter, 1999.
- [12] M. Schuld, I. Sinayskiy, and F. Petruccione, “The quest for a quantum neural network,” *Quantum Information Processing*, vol. 13, no. 11, pp. 2567–2586, 2014.
- [13] M. Altaisky, “Quantum neural network,” *arXiv preprint quant-ph/0107012*, 2001.
- [14] L. Fei and Z. Baoyu, “A study of quantum neural networks,” in *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on*, vol. 1, pp. 539–542, IEEE, 2003.
- [15] R. Zhou and Q. Ding, “Quantum mp neural network,” *International Journal of Theoretical Physics*, vol. 46, no. 12, pp. 3209–3215, 2007.
- [16] M. Siomau, “A quantum model for autonomous learning automata,” *Quantum information processing*, vol. 13, no. 5, pp. 1211–1221, 2014.
- [17] A. Sagheer and M. Zidan, “Autonomous quantum perceptron neural network,” *arXiv preprint arXiv:1312.4149*, 2013.
- [18] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: a new learning scheme of feed-forward neural networks,” in *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, vol. 2, pp. 985–990, IEEE, 2004.
- [19] E. C. Behrman, L. Nash, J. E. Steck, V. Chandrashekar, and S. R. Skinner, “Simulations of quantum neural networks,” *Information Sciences*, vol. 128, no. 3, pp. 257–269, 2000.
- [20] D. Ventura and T. Martinez, “An artificial neuron with quantum mechanical properties,” in *Artificial Neural Nets and Genetic Algorithms*, pp. 482–485, Springer, 1998.

- [21] R. Zhou, L. Qin, and N. Jiang, “Quantum perceptron network,” in *Artificial Neural Networks–ICANN 2006*, pp. 651–657, Springer, 2006.
- [22] A. Narayanan and T. Menneer, “Quantum artificial neural network architectures and components,” *Information Sciences*, vol. 128, no. 3, pp. 231–255, 2000.
- [23] M. Panella and G. Martinelli, “Neural networks with quantum architecture and quantum learning,” *International Journal of Circuit Theory and Applications*, vol. 39, no. 1, pp. 61–77, 2011.
- [24] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain,” *Psychological review*, vol. 65, no. 6, p. 386, 1958.